



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,872

DATE: 10/07/2002 TIME: 16:57:42

Input Set : A:\09010-108001.TXT

```
4 <110> APPLICANT: Callen, Walter
        Richardson, Toby
        Frey, Gerhard
 6
        Short, Jay M.
 7
        Mathur, Eric J.
         Gray, Kevin A.
 9
         Kerovuo, Janne S.
10
         Slupska, Malgorzata
11
13 <120> TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
        AND METHODS OF USE THEREOF
16 <130> FILE REFERENCE: 09010-108001
18 <140> CURRENT APPLICATION NUMBER: US 10/081,872
                                                           ENTERED
19 <141> CURRENT FILING DATE: 2002-02-21
21 <150> PRIOR APPLICATION NUMBER: US 60/270,495
22 <151> PRIOR FILING DATE: 2001-02-21
24 <150> PRIOR APPLICATION NUMBER: US 60/270,496
25 <151> PRIOR FILING DATE: 2001-02-21
27 <150> PRIOR APPLICATION NUMBER: US 60/291,122
28 <151> PRIOR FILING DATE: 2001-05-14
30 <160> NUMBER OF SEQ ID NOS: 321
32 <170> SOFTWARE: FastSEQ for Windows Version 4.0
34 <210> SEQ ID NO: 1
35 <211> LENGTH: 1311
36 <212> TYPE: DNA
37 <213> ORGANISM: Artificial Sequence
39 <220> FEATURE:
40 <223> OTHER INFORMATION: synthetically generated oligonucleotide
42 <400> SEQUENCE: 1
                                                                           60
43 atggcaaagt attccgagct cgaagagggc gggctcataa tgcaggcctt ctactgggac
44 gtccccatgg gaggaatctg gtgggacacg atagcccaga agatacccga ctgggcaagc
                                                                          120
45 gccgggattt cggcgatatg gattcccccg gcgagcaagg gcatgggcgg cgcctattcg
                                                                          180
46 atgggetacg acceetacga ettetttgae eteggtgagt acgaecagaa gggaacggta
                                                                          240
47 gagacgeget ttggetecaa geaggagete gtgaacatga taaacacege ecaegeetat
                                                                           300
48 ggcatgaagg taatagccga tatagtcatc aaccaccgcg ccggcggtga cctggagtgg
                                                                           360
49 aaccectteg tgaacgacta tacctggace gactteteaa aggtegegte gggtaaatae
                                                                           420
50 acggccaact acctcgactt ccacccgaac gagctccatg cgggcgattc cggaacattt
                                                                           480
                                                                           540
51 ggaggctatc ccgacatatg ccacgacaag agctgggacc agtactggct ctgggccagc
52 caggagaget acgeggeata teteaggage ateggeateg atgeetggeg ettegaetae
                                                                           600
53 gtcaagggct acggagcgtg ggtcgtcaag gactggctgg actggtgggg aggctgggcc
                                                                           660
54 gtcggggagt actgggacac aaacgttgat gcactgetca actgggccta ctcgagcgat
                                                                           720
55 gcaaaagtet tegaetteee getetactae aagatggaeg eggeetttga eaacaagaae
                                                                           780
                                                                           840
56 attoccgcac togtogaggo cotcaagaac gggggcacag togtoagoog cgaccogttt
57 aaggeegtaa eettegttge aaaceaegae aeegatataa tetggaacaa gtateeagee
                                                                           900
```





## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,872

DATE: 10/07/2002 TIME: 16:57:42

Input Set : A:\09010-108001.TXT

58 tacgcgttca tcctcaccta cgagggccag ccgacaatat tctaccgcga ctacgaggag 59 tggctcaaca aggataaget caagaacctc atctggatac atgacaacct cgccggagga 60 agcactgaca tcgtttacta cgacaacgac gagctgatat tcgtgagaaa cggctacgga 61 agcaagccgg gactgataac atacatcaac ctcgcctcaa gcaaagccgg aaggtgggtt 62 tacgttccga agttcgcagg ctcgtgcata cacgagtaca ccggcaatct cggcggctgg 63 gtggacaagt gggtggactc aagcggctgg gtctacctcg aggctcctgc ccacgacccg 64 gccaacggcc agtacggcta ctccgtctgg agctactgcg gtgttgggtg 66 <210> SEQ ID NO: 2 67 <211> LENGTH: 436 68 <212> TYPE: PRT 69 <213> ORGANISM: Artificial Sequence 71 <220> FEATURE:	960 1020 1080 1140 1200 1260 1311
72 <223> OTHER INFORMATION: synthetically generated polypeptide	
74 <400> SEQUENCE: 2 75 Met Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Leu Ile Met Gln Ala 10 15	
76 1 77 Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp Trp Asp Thr Ile Ala	
78 79 Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile	
81 Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp	
83 Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val	
85 Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr 90 95	
87 Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His	
89 Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr	
91 Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr	
93 Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe	
95 Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp	
96 97 Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly	
98 180 180 183 99 Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val	
100 195 200 101 Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr	
102 210 215 220 103 Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp	
104 225 230 230 233	
105 Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe	
106 107 Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly	
108 260 265 270 109 Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn	

PATENT APPLICATION: US/10/081,872 DATE: 10/07/200
TIME: 16:57:42

DATE: 10/07/2002

Input Set : A:\09010-108001.TXT

280 285	
110 275 260 111 His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile	
113 Low Thr Tyr Glu Gly Gln Pro Thr 11e Pne 191 Alg ASP 191 Can	
115 Trp Leu Asn Lys Asp Lys Leu Lys Ash Leu He He He He	
117 tow Ala Cly Cly Ser Thr Asp Ile Val Tyr Tyr Asp Ash Asp of a land	
110 The phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu IIe Thr Tyr	
121 The Asp Leu Ala Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pio Lys	
121 Tie Asii Bed Mid 555 380	
122 370 375 123 Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp	
124 385 125 Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro 125 Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro 410 415	
126 405 127 Ala His Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr	
128	
129 Cys Gly Val Gly	
130 435	
132 <210> SEQ ID NO: 3	
133 <211> LENGTH: 1311	
134 <212> TYPE: DNA	
135 <213> ORGANISM: Artificial Sequence	
137 <220> FEATURE: 138 <223> OTHER INFORMATION: synthetically generated oligonucleotide	
140 <400> SEQUENCE: 3 141 atggccaagt acctggaget cgaagaggge gggctcataa tgcaggcett ctactgggae	60
	120
	180
143 geogggattt eggegatatg gatteeeeg gegageday, staasja aggaaeggta 144 atgggetaeg acceptaega ettetttgae eteggtgagt acgaecagaa gggaaeggta	240
144 atgggetacg acceptacga ettertigae eteggegage dogada postace etaggeetac 145 gagacgeget ttggetecaa geaggagete gtgaacatga taaacacege eeaegeetac	300
145 gagacgcgct ttggctccaa gcaggagete gtgadcacgc ccggaggaga ccttgagtgg	360
145 gagacgeget tiggetecaa geaggagete gigadedega edalam 146 ggeateaagg teategeaga catagtaate aaceaeegeg eeggaggaga eetigagtgg 147 aaceeetteg teaatgaeta eacetggaeg gaettetega aggiegette eggaaeatt	420
147 aaccectteg teaatgacta caeeeggaeg gaetteetga caggegatte eggaacattt	480
147 aacccetteg teaatgacta caectygueg guetestyn by 148 acggeeaatt acctegactt ceaecegaac gageteeatg egggegatte eggaacattt 148 acggeeaatt acctegactt ceaecgacaag agetgggace agtaetgget etgggeeage 149 ggaggetate eegacatatg eeaeggacaag atggeateg atgeetggeg ettegactae	540
149 ggaggetate ecgacatatg ecaegacaag agetggggee agentggeg ettegactae	600
149 ggaggetate ecgacatatg ecaegacaag agetgggtoo agus 150 caggagaget acgeggeata teteaggage ateggetga actggtggg aggetgggeg	660
	720
	780
	840
	900
	960
	1020
	1080
	1140
	1200
	1260
160 tacgttccga agttcgcggg agcgtgcate cacgageded 1939 161 gtggacaagt gggtggactc aagcgggtgg gtgtacctcg aggcccctgc ccacgacccg	
<del>-</del> ·	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,872

DATE: 10/07/2002 TIME: 16:57:42

Input Set :  $A:\09010-108001.TXT$ 

162	62 gccaacggot attacggota ctccgtctgg agctattgcg gtgttgggtg a											1311					
164	4 <210> SEQ ID NO: 4																
	5 <211> LENGTH: 436																
	6 <212> TYPE: PRT																
167	<213	> OR	GANI	SM:	Arti	fici	al S	eque	nce								
169	<220	> FE	ATUR	E:													
170	<223	ro <	HER	INFO	RMAT	ON:	syn	thet	ical	.ly g	ener	rated	l pol	.ypep	otide	9	
172	<100	> SF	OHEN	CE:	4												
173	Met	Ala	Lys	Tyr	Leu	Glu	Leu	Glu	Glu	Gly	Gly	Leu	Ile	Met	GIn	Ala	
171	1				5					10					TO		
175	Phe	Tyr	Trp	Asp	Val	Pro	Met	Gly	Gly	Ile	Trp	Trp	Asp	Thr	тте	Ата	
176				20					25					30			
177	Gln	Lys	Ile	Pro	Asp	Trp	Ala	Ser	Ala	Gly	Ile	Ser	Ala	TTE	Trp	ire	
170			35					40					45				
179	Pro	Pro	Ala	Ser	Lys	Gly	Met	Gly	Gly	Ala	Tyr	Ser	Met	Gly	Tyr	Asp	
180		50					55					bυ					
181	Pro	Tyr	Asp	Phe	Phe	Asp	Leu	Gly	Glu	Tyr	Asp	Gln	Lys	GLY	Thr	vaı	
100	65					70					75					80	
183	Glu	Thr	Arg	Phe	Gly	Ser	Lys	Gln	Glu	Leu	Val	Asn	Met	iie	Asn	Thr	
19/					85					90					90		
185	Ala	His	Ala	Tyr	Gly	Ile	Lys	Val	Ile	Ala	Asp	IIe	Val	11e	Asn	HIS	
186				100					105			** 1	•	110	П	mb∽	
187	Arg	Ala	Gly	Gly	Asp	Leu	Glu	Trp	Asn	Pro	Phe	Val	Asn	Asp	туг	THE	
188			115					120			_	<b></b>	125	210	7 an	Шттг	
189	Trp	Thr	Asp	Phe	Ser	Lys	Val	Ala	Ser	GTĀ	Lys	Tyr	Thr	Ala	ASII	тут	
190		130					135		•	- 1	<b>a</b> 1	140	0.00	C1	mb x	Dho	
191	Leu	Asp	Phe	His	Pro	Asn	Glu	Leu	Hls	Ата	GLY	Asp	ser	СТА	1111	160	
192	145					150	_	•		T	155	M ~~~	N an	Cln	Пτιν		
193	Gly	Gly	Tyr	Pro		Ile	Cys	His	Asp	Lys	ser	ттр	ASP	GIII	175	115	
194					165		_	m	. 1 .	170	M	T 011	λνα	Sar		Glv	
	Leu	Trp	Ala			GLu	ser	Tyr	ALa	Ald	тут	ьeu	AIG	190	110	GII	
196			_	180	_	-1	•	m	185		C1**	TP T T T	λla		Tro	Va1	
	Ile	Asp		Trp	Arg	Pne	Asp	Tyr	vaı	гуѕ	GTĀ	тут	205	110	111	,	
198	_		195	_	<b>.</b>	3	m	200	C1,,	C1 17	Trn	Δla		Glv	Glu	Tvr	
				Trp	Leu	ASI	TIP	тър	GTA	GIY	115	220	Val	O <sub>1</sub>	014	Tyr	
200		210		_	** 7	<b>3</b>	215	17a l	T OU	A an	Trn		Tur	Ser	Ser	Glv	
		Asp	Thr	Asn	vaı			Val	ьeu	ASII	235	AIU	1 7 1	DCI	001	Gly 240	
202	225			_,	•	230	310	Tou	Marx.	Пттт			Δsn	Glu	Ala		
	Ala	Lys	Val	Pne	ASP	Pne	Ата	Leu	ı yı	250	цуз	1100	шьр	010	255		
204	_	_	<b>.</b>	3	245	Dwo	λ 1 a	TOU	Wa 1			Leu	Gln	Asn		Gln	
		Asn	ьуs			PLO	Ата	ьeu	265	261	AIU	ncu	0111	270	0-1		
206	_,	1	** . 1	260	. 7	1 an	Dro	Dho			Va 1	Thr	Phe			Asn	
		Val			Arg	ASP	PIO	280	цуз	пти	val	1111	285				
208	•	_	275	3	T1.	т1.	Пrn	200 Acn	Lvc	ጥህጉ	Pro	Ala			Phe	ıle	
				Asp	тте	ıre	295	MOII	пуз	- Y -	110	300	- 1 -			_	
210		290	m	<i>α</i> 1	C1	. Cln			T10	Pho	Tvr			Tvr	Glu	Glu	
			туr	GIU	. сту	310		1117	* T C		315	5	F	1 -		320	
212	305	T ~·-	<b>7</b> ~ ~	T tro	λαν	O I O	Τ,Δ11	Lve	Agn	Len			Ile	His	Asp	Asn	
213	rrp	ьeu	ASN	- БАЗ	, wah	, пуз	neu	. Lys	11011	u		F		_	-		

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,872 TIME: 16:57:42

DATE: 10/07/2002 TIME: 16:57:42

Input Set : A:\09010-108001.TXT

													225		
214			325					330					335	_	
215	Leu Ala Gly	Gly	Ser '	Thr	Asp	Ile		Tyr	Tyr	Asp	Asn	Asp	Glu	Leu	
216		340					345					350			
217	Ile Phe Val	Arg	Asn (	Gly	Tyr	Gly	Ser	Lys	Pro	Gly	Leu	Ile	Thr	Tyr	
218	355					360					365				
219	Ile Asn Leu	Ala	ser	Ser	Glu	Ala	Gly	Arg	Trp	Val	Tyr	Val	Pro	Lys	
220	370				375					380					
221	Phe Ala Gly	Ala	Cys	Ile	His	Glu	Tyr	Thr	Gly	Asn	Leu	Gly	Gly	Trp	
222	385			390					395					400	
223	Val Asp Lys	Trp	Val .	Asp	Ser	Ser	Gly	Trp	Val	Tyr	Leu	Glu	Ala	Pro	
224			405					410					415		
225	Ala His Asp	Pro	Ala .	Asn	Gly	Tyr	Tyr	Gly	Tyr	Ser	Val	Trp	Ser	Tyr	
226		420			_		425					430			
	Cys Gly Val	Gly													
228	435														
	<210> SEQ I		5												
	<211> LENGT														
	<212> TYPE:														
	<213> ORGAN		Arti	fic	ial :	Seque	ence							•	
	<220> FEATU					-									
236	<223> OTHER	INFO	RMAT	ION	: sv	nthe	tical	11y (	gene:	rate	l ol	igon	ucle	otide	
	<400> SEQUE				4			-	_						
230	atggccaagt	actcc	egage	t. ac	raag	aggg	ट वव	catt	ataa	tgc	aggc	ctt	ctac	tgggac	60
240	gtcccaggtg	ааааа	atct	ar art	Laga	acac	c at	caqq	agca	aga:	tacc	gga	gtgg	tacgag	120
240	gcgggaatat	CCGCC	attt	ar ara	atte	cccc	a ac	aaqc	aaqq	qca	tagg	cgg	cgcc	tattcg	180
241	atgggctacg	acccc	rtacq	a ci	ttat	t.t.ga	c ct	caat	gagt	acq	acca	qaa	ggga	acggta	240
242	gagacgcgct	ttaac	rtcca	a o	cado	aget	c at	gaac	atqa	taa	acac	cgc	ccac	gcctat	300
243	ggcatgaagg	taata	acca	ıa ta	atag:	tcat	c aa	ccac	caca	ccq	acaa	tga	cctq	gagtgg	360
244	aaccccttcg	tgaac	rgact	a ta	acct	ggac	c αa	at.t.a	tcaa	agg	tcac	qtc	gggt	aaatac	420
243	acggccaact	accto	rgact	t 0	cacc	caaa	e ga	acto	cato	caa	acaa	ttc	caga	acattt	480
240	ggaggctatc	accec	ratat	or co	caco	acaa	or aon	ctaa	gacc	aαt.	acto	act	ctaa	gccagc	540
247	caggagagct	20000	racat	a t	ctca	асаа	c at	caac	atco	at.g	cata	aca	cttc	gactac	600
240	gtcaagggct	atrot	ccat	ים מי	atca	tcaa	a da	ctaa	ctaa	act	aata	aaa	aggc	tagaca	660
249	gttggagagt	acyce	.0000	9 9	aaca	toga	ישע כימרי	tatt	ctca	act	aaac	at.a	ctca	agcggt	720
250	gccaaggtct	ttase	yyaca sttaa		atet	acta	c aa	αatα	gatg	agg	cett	t.ga	caac	aaaaac	780
25T	attccagcgc	tagte	2 t a t a		atta	2022	c aa	ccan	acto	++a	tete	aga -	cgac	ccattc	840
252	actecagege	costt	eretar		2200	agaa	c gg	coat	ataa	tet	nnaa	caa	atac	cttact	900
253	tatgctttca	CCLLI	LgLag	ic a	aacc	acya	a aa	cata	atat	tet	acca	caa	ccac	gaggag	960
254	tggctcaaca	teete	Jacci	.a c	yaay	goda agat	y cc	ey cc	atac	aca	acca	cct	caca	aataaa	1020
255	tggctcaaca	aggac	caggu	L g	aaca	acct	t at	acyy	atat	tog	tcan	gaa	caac	tacada	1080
256	agcaccgaca	tagto	ctact	ac	yala	acya +aaa	t ya	2000	4001	aca	agge	caa	aadd	taaatt	1140
257	gacaagccgg	ggcti	cacaa	ic c	Laca	ccaa	0 01	ayyc	toya tata	gea	ayyc	cet	caaa	aactaa	1200
258	tatgtgccga	agtto	egegg	ig c	gcgı	gcat	- ~+	cyay	ataa	220	g caa ataa	200	ttac	gaccet	1260
259	gtagacaagt	acgto	ctact	.c a	aycg	yctg +a+a	y yc	otac	+444	aay	taaa	ayu	a	940000	1311
	gccaacgggc			.a c	LCCG	Lytg	y ag	CLac	Lycy	999	-999	cly	u		1911
	<210> SEQ I														
	<211> LENGT		36												
	<212> TYPE:					a									
	<213> ORGAN		Arti	LIIC	ıaı	sequ	ence								
267	<220> FEATU	JRE:													

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/081,872

DATE: 10/07/2002 TIME: 16:57:43

Input Set : A:\09010-108001.TXT